**Title:** A system and method for fingerprint based media recognition

**Inventors:**

Sean Ward (14 West Cedar Street, Alexandria, Va 22301)

Isaac Richards (2959 Reeves Road, Willoughby, OH 44094)

**[0001]   Related Applications:**

This application claims the benefit of the filing date of provisional application

60/454,329, filed March 14, 2003, and titled "A System And Method For Fingerprint

Based Media Recognition".

**[0002]   Field of the Invention:**

The present invention is related to a method for the recognition of media, such as

CDs or DVDs. More specifically, it relates to the recognition of media using a

combination of acoustic and bit based fingerprints, and segment length information.


**[0003]   Description of the Prior Art:**

Generally, media identification has been based on either the recovery of specially

formatted metadata fields within the media, such as CD-TEXT in CDs, or on identifying

identical pressings of a mass produced piece of media, such as CD table-of-contents

information. Examples of table-of-contents based systems include US6,061,680, used in

the commercial CDDB system by Gracenote, and the Musicbrainz and FreeDB systems

available from open source public systems.

**[0004]**   To address the limitations of TOC based systems, fingerprint based systems are

able to identify items on a track level basis without embedding information. Examples of

acoustic fingerprinting systems include US2002161741, US20020133499, and

US20020083060. These systems however are unable to leverage the fact that most media

is still mass produced, which allows additional pieces of information to aid in the

identification of said media. Finally, such systems are unable to recognize media with

pure data segments, such as computer data CDs.

**[0005]**   Finally, bit based solutions (www.bitzi.com) have attempted to address the

issue of file or media identification. These rely upon the computation of a bit-based hash,

such as an MD5sum, or a tigertree hash, which determines how identical two files or

media segments are. However, such systems are unable to cope with user created content, such as burned CDs, or format shifted media.

**[0006]    Summary of the Invention:**

This system for media recognition comprises two major parts: the media analysis component, and the media recognition component. Table of contents information (consisting of a table indicating the number and length of segments contained on the media) and an acoustic or bit based fingerprint of the contents of one or more segments from the media is collected by the media analysis component. This information is then used by the media recognition component to identify the media, and in the case that no matching media record is found, acoustic or bit-based fingerprints can be extracted from the remaining segments to attempt partial recognition on a per segment basis.

[0007]    It is therefore an object of this invention to allow the recognition of both commercially available and user created media, in situations where existing segment length analysis fails. It is also an object of this invention to allow the partial identification of new media, when it contains any segments that existed on existing, indexed media. Additionally, it is an object of this invention to provide a useful balance between accuracy and computation cost of recognition, which a system built purely on acoustic fingerprinting, fails to achieve in the context of strictly media recognition. Finally, it is an object of this invention to provide accurate identifications of media with low segment counts, which have a poor accuracy rate in a pure segment length analysis.

**[0008]    Brief Description of the Drawings:**

In the drawings:

[0009]    FIG. 1 is a logic flow diagram, showing the overview process of fingerprint-based media recognition.

[0010]    FIG. 2 is a block diagram, showing the components of the media recognition component.

[0011]    FIG. 3 is a logic flow diagram, showing the process of recognizing a piece of media from the summary fingerprints and table of contents information.

**[0012]   Detailed Description of the Invention:**

The ideal context of this system places the media analysis component within a media playback tool, such as a software media player or a hardware CD player. Referring to the flow diagram of Fig. 1, this system, upon a new piece of media, such as a CD or DVD, being inserted at access media step 10, proceeds to extract the table of contents segment information in step 20, and, depending on whether the segments within the media are data or audio, fingerprint one or more segments (step 30) to derive a media description packet. This media description packet is then transmitted to the media recognition component (step 40) for resolving the media identification request, the identification using the process illustrated in the flow diagram of FIG 3.

**[0013]**   Ideally, the media recognition component (FIG 2) is located on a remote server, using TCP/IP or http for communications. This allows a large-scale database to be centrally managed without replicating the database on each media identification client. However, in certain embedded applications, such as media player hardware units, which lack connectivity, the media recognition component may exist on the same device as the media analysis component.

**[0014]**   The first step of recognition is the resolution of the fingerprints in the media description packet (step 120) wherein one or more track fingerprints are received. Depending on the type of fingerprint, this may require a query (step 130) against a reference acoustic fingerprint database (100), or reference bitprint database to resolve the fingerprint identification. In the context of a hash bitprint, the print may be the fingerprint identifier, such as with an MD5 sum. A query for table of content records containing the fingerprint identifiers and segment count in the incoming media description packet is then performed (step 140) using the TOC mapping database (90). Finally, that result set is culled based on the segment lengths matching those within the incoming media description packet.

**[0015]**   In the event that the resulting media description record set contains more than one entry, or is empty, a response is sent back to the media analysis component requesting the fingerprints for all remaining media segments (step 60 and step 160). This allows the system to fall back to a segment level identification for user created media,

3

such as burned CD's. Upon receiving the full set of fingerprints from the media analysis component (step 70), the recognition component resolves the fingerprint identifiers for each segment (step 170) using the fingerprint database (100). If all segments within the media matched known fingerprints, a new media description record can be automatically added to the system at this point as well.

[0016]     In the event that the media description record set contains only one entry, then the fingerprint identifiers for all un-fingerprinted segments in the media can be retrieved from the description record, saving the cost of fingerprinting and resolving each segment individually (step 180).

[0017]     The final step in the recognition process is the retrieval of the appropriate metadata for the media, using the segment level fingerprint identifiers and potentially a media identification identifier (step 190). This allows the returned metadata to account for duplicated segments on different media, such as returning the appropriate album for an audio track that appears on multiple CD's, and is stored in the identifier to metadata mapping database (110).

[0018]     In the case where no fingerprint segments, or media description records match an incoming media description packet, a request can be sent back to the media analysis component that the user manually identify the work. This allows the system to index new media as it is encountered in actual usage. The manually identified media description record can then be sent from the media analysis component to the central media recognition component, where it can be stored for later addition to the system. Many insertion strategies are possible, including requiring a threshold of similar descriptions for a new media entry be collected before insertion occurs, or that human review is needed to allow the new entry to be added to the database.

[0019]     While this invention has been described in conjunction with specific embodiments thereof, it is evident that many alternative modifications and variations will be apparent to those skilled in the art. Accordingly, the preferred embodiments of the invention as set forth herein are intended to be illustrative, not limiting. Various changes may be made without departing from the true spirit and scope of the invention as defined in the following claims.